

文章编号: 0253-3782(2003)04-0382-08

计算台站接收函数的最大熵谱反褶积方法^{*}

吴庆举 田小波 张乃铃 李卫平 曾融生

(中国北京 100081 中国地震局地球物理研究所)

摘要 提出一种在时间域采用最大熵谱反褶积提取台站接收函数的方法, 以最大熵作为自相关函数和互相关函数的递推准则, 利用 Toeplitz 方程及 Levinson 递推算法, 得到预测误差滤波系数的递推公式, 从而计算台站接收函数. 外推运算过程中, 反射系数总是小于 1, 保证了最大熵谱反褶积的稳定性. 时窗外数据熵极大提高了接收函数的分辨率. 合成地震图与实测地震图的检验表明, 最大熵谱反褶积是一种在时间域测定台站接收函数的有效方法.

关键词 接收函数 最大熵谱 Toeplitz 方程 Levinson 递推算法

中图分类号: P315.0 **文献标识码:** A

引言

接收函数是用远震 P 波波形的垂直分量对径向分量和切向分量作反褶积处理后得到的时间序列. 它是台站下方地壳上地幔速度结构远震平面波的响应, 基本上与震源及传播路径无关. 由于实际的远震资料是有限带宽的, 且含有各种噪声, 垂直分量一般都存在近零的频率成分, 导致频率域反褶积不够稳定, 故要对垂直分量作预白化处理, 以压制其近零值的频谱成分 (Owens *et al.*, 1987; 刘启元等, 1996).

从三分量远震 P 波波形中分离出可靠的台站接收函数, 对用远震 P 波波形反演接收介质结构起着决定性的作用. 接收函数一般是通过远震 P 波波形的水平与垂直分量的频谱相除, 然后反变换到时间域而得到的. 由于垂直分量常含有近零值振幅, 频率域除法运算一般不够稳定, 常靠引入“水准量”来保证频率域除法运算的稳定, 但降低了接收函数的精度; 此外, 频率域运算不仅对数据长度有严格要求, 而且隐含着时窗之外数据为零或呈周期变化的不合理假设, 数据长度及采样间隔限制了频率域反褶积的分辨率. 为了克服频率域反褶积的缺点, 本研究提出一种在时间域用最大熵谱反褶积计算台站接收函数的方法.

最大熵谱反褶积以熵极大作为自相关函数和互相关函数的外推准则, 来求取接收函数. 在外推运算过程中, 反射系数总是小于 1, 从而保证了最大熵谱反褶积的稳定性. 时窗外数据最不可预测性增强了接收函数的分辨率.

时间序列分析中, 要获得精确的谱, 则要求时间序列必须足够长, 足以包含最低有效频率成分对时间序列的贡献. 在对实际有限长时间序列作谱分析时, 自相关方法和傅氏变

^{*} 国家自然科学基金(49974021)资助项目. 中国地震局地球物理研究所论著 03AC1021.
2002-07-26 收到初稿, 2003-02-24 收到最后修改稿并决定采用.

换法假设数据时窗之外的数据要么为零, 要么呈周期性变化, 这种假设显然与实际相悖. 如果能够突破常规方法, 对数据时窗之外的数据所做的不符合客观实际的假设, 引入更加合理的假设, 则必定能够进一步提高谱的分辨率, 获得高分辨率的频谱.

Burg(1972)提出了最大熵谱法, 旨在提高频谱的分辨率. 其基本思想是: 假设时窗外的数据具有最不可预测性, 在每次外推自相关函数时, 都要使其相应的熵极大化. Van den Bos(1971)指出, 最大熵谱法等价于自回归模型的最小二乘解. 自回归模型在工程上称为全极点模型(Ulrych, Bishop, 1975; Ulrchy, Clayton, 1976).

对 Wiener 滤波加以改造, 以熵极大作为自相关函数和互相关函数的外推准则, 可以得到最大熵谱反褶积. 最大熵谱反褶积已广泛用于地震勘探数据处理, 用于压缩子波, 提取反射系数序列, 提高分辨率(Clarerbout, 1976). Tselentis(1990)曾引入最大熵谱反褶积, 计算台站间格林函数. 本研究将对最大熵谱反褶积加以发展, 用于提取台站接收函数, 以获得高分辨率的接收函数.

1 最大熵谱的概念

设有概率空间 X

$$X = \{x_i, p_i\} \quad i = 1, 2, \dots, N \quad (1)$$

式中, x_i 是该空间可能发生的事件, p_i 为其发生的概率. 如果已知某一事件发生的概率, 就获得了关于该概率空间的部分信息, 信息量与概率之间的关系可表示为

$$I = k \ln \frac{1}{p_i} \quad (2)$$

其中, k 是一个常数, 当对数以 2 为底时, $k=1$.

对于确定事件, $p_i=1$, 故信息量 $I=0$. 对于随机事件, 信息量为一正数.

如果对上述概率空间作长时间的观测, 在长度为 t 的时间间隔内, 各事件发生的次数为 $P_i T$, 于是总的信息量为

$$I_{\text{total}} = k \left(p_1 t \ln \frac{1}{p_1} + p_2 t \ln \frac{1}{p_2} + \dots \right) \quad (3)$$

所谓熵是指概率空间单位时间内的平均信息量, 即

$$H = \frac{I_{\text{total}}}{t} = -k \sum_{i=1}^M p_i \ln p_i \quad (4)$$

由信息量与概率之间的关系可知, 熵是描述概率空间不确定性的一种测度. 如果概率空间某一事件的发生概率为 1, 则其余事件的发生概率为 0, $H=0$, 这是一个完全确定的系统; 除此之外, 熵只可能是一正数; 当一系统具有最不确定性时, 其熵达到最大.

平稳高斯过程 x_i 的功率谱 $S(f)$ 与熵的关系可表示为

$$H = \frac{1}{4f_N} \int_{-f_N}^{f_N} \lg S(f) df \quad (5)$$

其中, f_N 是 Nyquist 频率. 式(5)可用自相关函数 $\phi(k)$ 表示为

$$H = \frac{1}{4f_N} \int_{f_N}^{f_N} \lg \left[\sum_{-\infty}^{+\infty} \phi(k) \exp(-i2\pi f k \Delta t) \right] df \quad (6)$$

其中, Δt 是采样间隔.

假设自相关函数 $\phi(1), \phi(2), \dots, \phi(m-1)$ 为已知, 以这些已知自相关函数作为约束条

件, 对式(6)求关于待定的自相关系数的极大, 可求取未知的 $\phi(m), \phi(m+1), \dots, \phi(N)$. 这种以熵达极大来外推自相关函数而得到的功率谱, 就是所谓的最大熵谱, 可表示为

$$P_E(f) = \frac{P_m}{f_N \left| 1 + \sum_{j=1}^{M-1} \gamma_j \exp(-i2\pi f j \Delta t) \right|^2} \quad (7)$$

其中, P_m 是一个常数, γ_j 是由数据 x_t 中得到的预测误差系数.

2 自回归过程与最大熵谱

离散的正态随机过程可用一线性滤波系统表示为

$$x_t = \mu + a_t + \phi_1 a_{t-1} + \phi_2 a_{t-2} + \dots \quad (8)$$

其中, 时间序列 a_t 是一白噪序列, 其均值为零, 方差为 σ_a^2 ; μ 为 x_t 的均值.

假设线性滤波器可表示为

$$\Psi(z) = 1 + \phi_1 z + \phi_2 z^2 + \dots \quad (9)$$

其中, z 是单位延迟算子.

若 $\Psi(z)$ 稳定, 则 x_t 为平稳的随机过程. 假设已经得到了 μ 的估计, 并从时间序列 x_t 中去掉了 μ , 则 x_t 是均值为零的随机过程.

对式(8)作 z 变换, 得

$$X(z) = \Psi(z)A(z) \quad (10)$$

设 $\Phi(z) = \Psi^{-1}(z)$, 则有

$$\Phi(z)X(z) = A(z) \quad (11)$$

将式(11)反变换回时间域得

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + a_t \quad (12)$$

在推导式(12)过程中, 假设该线性过程可逆, 可逆的条件是 $\Phi(z)$ 在单位圆上或单位圆内收敛.

虽然式(12)是一无穷级数, 但一般来说, 用其中有限的若干项可以很好地描述这一随机过程, 于是有

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_M x_{t-M} + a_t \quad (13)$$

式(13)代表 M 阶的自回归过程. M 的有限性保证了自回归过程的可逆性, 若使其稳定, 则多项式 $1 - \alpha_1 z - \dots - \alpha_M z^M$ 的根必须位于单位圆外.

对式(13)两端同乘 x_{t-k} 并取数学期望, 当 $k > 0$ 时, $E(x_{t-k} a_t) = 0$, 故有

$$r(k) = \alpha_1 r(k-1) + \alpha_2 r(k-2) + \dots + \alpha_M r(k-M) \quad k > 0 \quad (14)$$

将 $k=1, 2, \dots, M+1$ 代入式(14), 得到 Yule-Walker 方程组

$$\begin{cases} r(1) - \alpha_1 r(0) - \dots - \alpha_M r(M-1) = 0 \\ r(2) - \alpha_1 r(1) - \dots - \alpha_M r(M-2) = 0 \\ \vdots \\ r(M+1) - \alpha_1 r(M) - \dots - \alpha_M r(1) = 0 \end{cases} \quad (15)$$

假设自协方差函数 $r(0), r(1), \dots, r(M)$ 为已知, 将它们代入式(15)中前 M 个方程中, 可求得系数 $r(M+1), r(M+2), \dots, r(N)$, 待定的 $r(M+1)$ 由下式决定:

$$\det \begin{bmatrix} r(1) & r(0) & \cdots & r(M-1) \\ r(2) & r(1) & \cdots & r(M-2) \\ \vdots & \vdots & \vdots & \vdots \\ r(M+1) & r(M) & \cdots & r(1) \end{bmatrix} = 0 \quad (16)$$

与式(5)完全等价的平稳高斯过程 x_t 的熵的另一种表示形式为

$$H = \frac{1}{2} \lg \{ \det [\mathbf{C}(\mathbf{M})] \} \quad (17)$$

其中, $\mathbf{C}(\mathbf{M})$ 是随机过程 x_t 的自协方差矩阵, 即

$$\mathbf{C}(\mathbf{M}) = \begin{bmatrix} r(0) & r(1) & \cdots & r(M) \\ r(1) & r(0) & \cdots & r(M-1) \\ \vdots & \vdots & \vdots & \vdots \\ r(M) & r(M-1) & \cdots & r(0) \end{bmatrix} \quad (18)$$

假设已知自协方差函数 $r(0), r(1), \cdots, r(M)$ 的精确值, 在每次外推自协方差系数 $r(M+1), r(M+2), \cdots, r(N)$ 等时, 以随机过程的熵达到极大作为准则, 可得到最大熵谱.

对 $\det[\mathbf{C}(M+1)]$ 求关于 $r(M+1)$ 的极大值, 得

$$\det \begin{bmatrix} r(1) & r(0) & \cdots & r(M-1) \\ r(2) & r(1) & \cdots & r(M-2) \\ \vdots & \vdots & \vdots & \vdots \\ r(M+1) & r(M) & \cdots & r(1) \end{bmatrix} = 0 \quad (19)$$

比较式(16)和(19)可知, 最大熵谱分析与随机过程的自回归模型是完全等价的. 对式(13)作 z 变换可求得自回归过程的谱, 即

$$X(z) - X(z)(\alpha_1 z + \alpha_2 z^2 + \cdots + \alpha_M z^M) = A(z) \quad (20)$$

于是有

$$|X(z)|^2 = \frac{|A(z)|^2}{|1 - \alpha_1 z + \alpha_2 z^2 + \cdots + \alpha_M z^M|^2} \quad (21)$$

将 $z = \exp(-i2\pi f)$ 代入式(21), 就可求得自回归过程的功率谱

$$S(f) = \frac{2\sigma_a^2}{\left| 1 - \sum_{j=1}^M \alpha_j \exp(-i2\pi f j) \right|^2} \quad (22)$$

比较式(7)和(22), 不难发现两者的等价性.

由式(13)可以明显地看出, 所谓自回归过程就是根据该过程前 $(M-1)$ 个值来预测当前值, 预测误差为 $p_m, 1, -\alpha_1, -\alpha_2, \cdots, -\alpha_m$ 为 M 阶预测滤波器系数.

在式(13)中, 用 $\gamma_1, \gamma_2, \cdots, \gamma_M$ 代替 $-\alpha_1, -\alpha_2, \cdots, -\alpha_M$, 即可得到预测误差滤波器. 当用时间序列的当前值和过去值作为输入时, 输出为一白噪, 即

$$a_t = x_t - \gamma_1 x_{t-1} - \gamma_2 x_{t-2} - \cdots - \gamma_M x_{t-M} \quad (23)$$

预测误差滤波因子的求解, 可归结为 Toeplitz 方程组的求解, 即

$$\begin{bmatrix} \hat{r}(0) & \hat{r}(1) & \cdots & \hat{r}(M-1) \\ \hat{r}(1) & \hat{r}(0) & \cdots & \hat{r}(M-2) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{r}(M-1) & \hat{r}(M-2) & \cdots & \hat{r}(0) \end{bmatrix} \begin{bmatrix} \hat{\alpha}_{M1} \\ \hat{\alpha}_{M2} \\ \vdots \\ \hat{\alpha}_{MM} \end{bmatrix} = \begin{bmatrix} \hat{r}(1) \\ \hat{r}(2) \\ \vdots \\ \hat{r}(M) \end{bmatrix} \quad (24)$$

自协方差系数可以用下述的常规方法计算：

$$\hat{r}(k) = \frac{1}{N} \sum_{t=1}^{N-|K|} (x_{t+k} - m)(x_t - m) \tag{25}$$

其中

$$m = \frac{1}{N} \sum_{t=1}^N x_t \tag{26}$$

3 最大熵谱反褶积

由公式(7)或(22)计算自回归过程的最大熵谱，要求预测滤波系数为已知，通过对 Toeplitz 方程组(24)的求解，固然可以求得自回归系数，但需要自协方差函数的先验估计，当数据时窗较短时，由式(24)、(25)和(26)解得的自相关函数精度较低。基于自回归过程的预测误差，Burg(1972)指出，将预测误差滤波器对数据分别作正向和反向的预测滤波，得到正向预测误差和反向预测误差，对两者的平方和求关于 M 阶自回归系数的极小，可以求得 M 阶自回归系数。

M 阶正向预测误差估计可表示为

$$e_M^f[n] = x_t[n] + \sum_{m=1}^M x_t[n-m] a_M[m] \tag{27}$$

M 阶反向预测误差估计可表示为

$$e_M^b(n-M) = x_t(n-M) + \sum_{m=1}^M x_t[n+m-M] a_M[m] \tag{28}$$

根据 Toeplitz 方程组及 Levinson 递推算法 (Clearbout, 1976) 可知，预测误差滤波系数存在如下的递推关系：

$$a_M[m] = a_{M-1}[m] + K_{M-1} a_{M-1}[M-m] \tag{29}$$

其中， $a_M[m]$ ， $a_{M-1}[m]$ 分别代表自回归过程 M 阶与 $M-1$ 阶预测误差系数； k_{M-1} 称为反射系数。

将式(29)分别代入到式(27)和(28)，得到关于正向与反向预测误差的递推关系式

$$\begin{cases} e_M^f[n] = e_{M-1}^f[n] + K_{M-1} e_{M-1}^b[n-M] \\ e_M^b[n-M] = e_{M-1}^b[n-M] + K_{M-1} e_{M-1}^f[n] \end{cases} \tag{30}$$

在给定的时窗范围内，对正向预测误差和反向预测误差的平方和求极小，也就是使熵达极大

$$q_M^f = \sum_{n=M}^N \{ [e_M^f[n]]^2 + [e_M^b[n-M]]^2 \} = \min \tag{31}$$

令

$$\frac{\partial q_M^f}{\partial k_{M-1}} = 0 \tag{32}$$

可求出反射系数。式(31)的求和运算只涉及时窗之内的数据，而与时窗之外的数据毫无关系。换言之，时窗之外的数据具最不可预测性。于是可得到

$$K_{M-1} = \frac{-2 \sum_{n=M}^N e_{M-1}^f[n] e_{M-1}^b[n-M]}{\sum_{n=M}^N [e_{M-1}^b[n-M]]^2 + [e_{M-1}^f[n]]^2} \tag{33}$$

显而易见， $K_{m-1} \leq 1$ ，这也是称之为反射系数的原因。

得到反射系数的表达式后, 就可以从 $(M-1)$ 阶的预测滤波因子和预测误差, 递推出 M 阶的预测滤波因子和预测误差.

根据上述求解预测误差系数的思路, 完全可以平行地推导出计算接收函数的递推公式.

用 $h_M(n)$ 表示待求的 M 阶接收函数, $x(n)$ 和 $y(n)$ 分别代表远震 P 波波形记录的垂直分量和水平分量, 根据 Wiener 滤波原理, 有

$$s(n) = y(n) - \sum_{m=1}^N x(n-m)h_M(m) \quad (34)$$

其中, $s(n)$ 为 Wiener 滤波的误差向量.

接收函数 $h_M(n)$ 可以通过如下的矩阵运算得到:

$$RH = C \quad (35)$$

其中, R 是远震 P 波波形记录垂直分量的自相关系数矩阵, C 是垂直分量与水平分量的互相关系数向量.

比较 $(M-1)$ 阶与 M 阶的矩阵方程 (35), 可以得到关于接收函数的递推公式

$$h_M = h_{M-1}[m] + L_{M-1}a_M(M-m) \quad (36)$$

将式 (36) 代入 (34), 得

$$s_M[n] = s_{M-1}[n] - L_{M-1}e_M^b[n-M] \quad (37)$$

在数据时窗范围内, 对 $s_M[n]$ 的平方和求关于 L_{M-1} 的极小, 可得

$$L_{M-1} = \frac{\sum_{n=M}^N s_{M-1}[n]e_M^b[n-M]}{\sum_{n=M}^N e_M^b[n]^2} \quad (38)$$

于是可逐步递推出台站接收函数, 每次外推时, 以熵达极大作为准则. 其初始条件为

$$\begin{cases} R[0] = \sum_{n=0}^N x_i^2[n], C[0] = \sum_{n=0}^N x_i[n]y_i[n], a_0[0] = 1 \\ h_0[0] = \frac{C[0]}{R[0]} \\ e^f[n] = e^b[n] = x_i[n] \\ S_0[n] = y_i[n] - h_0[0]x_i[n], n = 0, 1, \dots, N \end{cases} \quad (39)$$

4 资料检验

我们分别用合成地震图和实际观测到的远震资料, 对本研究所提出最大熵谱反褶积方法进行数值检验, 以考察并阐述最大熵谱反褶积的实际应用效果.

图 1 是采用“反射率法”得到的合成地震图及其合成接收函数. 图 2 是用一个实际观测到的综合了震源和路径效应的时间函数, 对图 1 所示的合成地震图进行褶积运算后所得到的虚拟观测地震图. 图 3 是对图 2 所示的虚拟观测地震图分别用频率域反褶积和最大熵谱反褶积得到的接收函数. 可以发现, 最大熵谱反褶积能得到与理论值基本相同的接收函数, 表明最大熵谱反褶积的有效性.

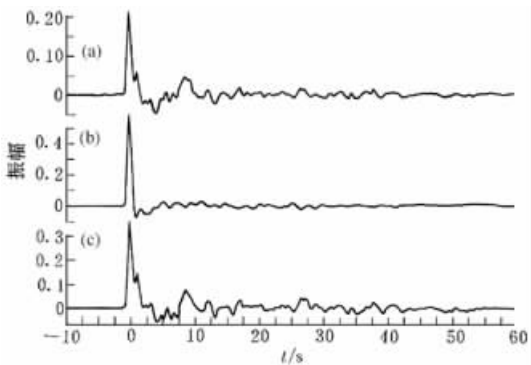


图 1 合成地震图及其合成接收函数
(a) 径向分量; (b) 垂直分量; (c) 合成接收函数

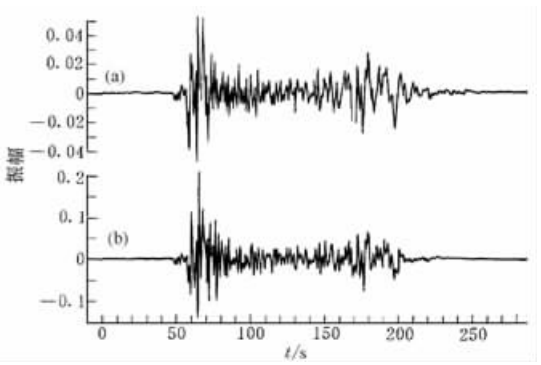


图 2 虚拟观测地震图(用一实测远震的垂直分量分别与图 1 的径向和垂直分量褶积得到)
(a) 径向分量; (b) 垂直分量

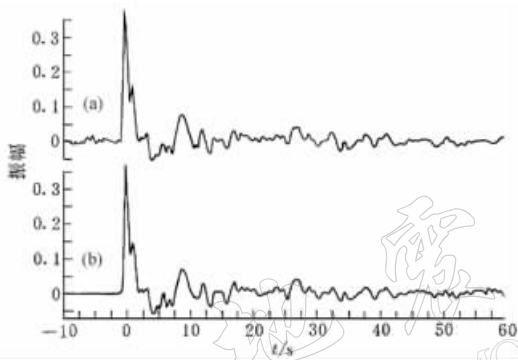


图 3 虚拟观测地震图的频率域反褶积与最大熵谱反褶积结果的比较
(a) 频率域反褶积; (b) 最大熵谱反褶积

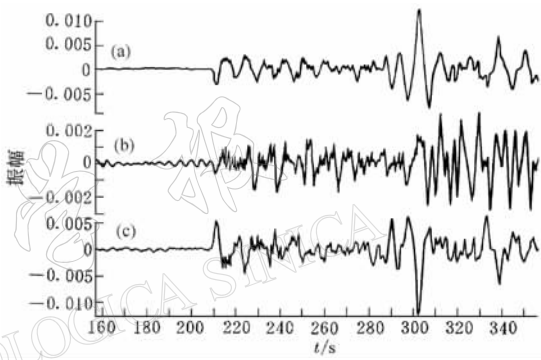


图 4 实测远震事件的三分量 P 波波形数据
(a) 径向分量; (b) 切向分量; (c) 垂直分量

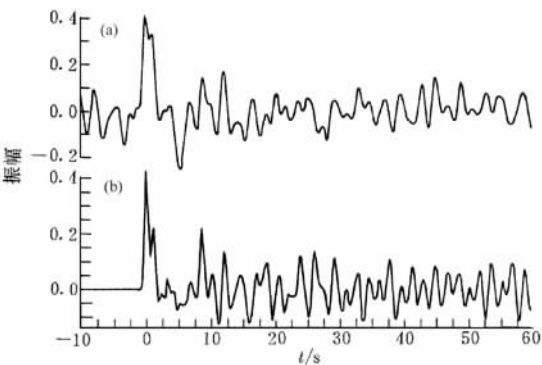


图 5 实测远震事件的频率域反褶积与最大熵谱反褶积结果的比较
(a) 频率域反褶积; (b) 最大熵谱反褶积

图 4 是某一远震事件的三分量 P 波波形数据. 图 5 是对图 4 所示的实测地震图分别用频率域反褶积和最大熵谱反褶积得到的接收函数. 从图 5 可以看出, 与频率域反褶积相比较, 最大熵谱反褶积所提取的接收函数的波形更加尖锐, 频率域的单峰值波形被分解成多峰值, 有效地提高了接收函数的分辨率. 此外, 最大熵谱反褶积能够有效地提取淹没于噪音中的受频率域反褶积水准量压制的有效信号.

合成地震图与实测地震图的检验表明, 最大熵谱反褶积是一种非常有效的在时间域提取接收函数的方法.

5 结论

最大熵谱反褶积在计算自相关函数和互相关函数时, 以极大熵作为外推准则, 避免了时窗外数据的主观假设. 由 Toeplitz 方程及 Levinson 递推算法得到的预测误差滤波系数的递推公式, 提高了接收函数的计算效率. 在外推运算过程中, 反射系数总是小于 1, 保证了最大熵谱反褶积的稳定性. 时窗外数据最不可预测性提高了接收函数的精度和分辨率. 合成地震图与实测地震图的检验表明, 最大熵谱反褶积能够有效地提取淹没于噪音中的有效信号, 并能从单峰值波形中分解出多峰值, 是一种在时间域测定台站接收函数的有效方法.

参 考 文 献

- 刘启元, Kind R, 李顺成. 1996. 接收函数复谱比的最大或然性估计及非线性反演[J]. 地球物理学报, **39**(4): 502~513
- Burg J P. 1972. The relationship between maximum entropy spectra and maximum likelihood spectra[J]. *Geophysics*, **37**(2): 375~376
- Clearbout J F. 1976. *Fundamentals of Geophysical Data Processing*[M]. New York: McGraw-Hill, 1~357
- Owens T J, Zandt G, Talor S R. 1987. Seismic evidence for an ancient rift beneath the Cumberland Plateau, Tennessee: A detailed analysis of broadband teleseismic P waveforms[J]. *J Geophys Res*, **89**(B9): 7 783~7 795
- Tselentis G A. 1990. Interstaion surface wave attenuation by autoregressive deconvolution[J]. *Pure Appl Geophys*, **133**(3): 429~446
- Ulrych T J, Bishop T N. 1975. Maximum entropy spectral analysis and autoregressive decomposition[J]. *Rev Geophys Space Phys*, **13**(1): 183~200
- Ulrych T J, Clayton R W. 1976. Time series modeling and maximum entropy[J]. *Phys Earth Plane Inter*, **12**(2/3): 188~200
- Van den Bos A. 1971. Alternative interpretation of maximum entropy spectral analysis[J]. *IEEE Trans Inform Theory*, **17**: 493~494

RECEIVER FUNCTION ESTIMATED BY MAXIMUM ENTROPY DECONVOLUTION

Wu Qingju Tian Xiaobo Zhang Nailong Li Weiping Zeng Rongsheng

(Institute of Geophysics, China Seismological Bureau, Beijing 100081)

Abstract: Maximum entropy deconvolution is presented to estimate receiver function, with the maximum entropy as the rule to determine auto-correlation and cross-correlation functions. The Toeplitz equation and Levinson algorithm are used to calculate the iterative formula of error-predicting filter, and receiver function is then estimated. During extrapolation, reflective coefficient is always less than 1, which keeps maximum entropy deconvolution stable. The maximum entropy of the data outside window increases the resolution of receiver function. Both synthetic and real seismograms show that maximum entropy deconvolution is an effective method to measure receiver function in time-domain.

Key words: receiver function; maximum entropy; Toeplitz equation; Levinson algorithm